基于多模态特征挖掘的新媒体版权保护方法研究与应用

王梦环 罗小龙 李梓华
（广东南方智媒科技有限公司，广东 广州 510000）

摘要：【目的】《知识产权强国建设纲要（2021—2035 年）》中将知识产权战略视为国家重要战略，强调知识产权在国家层面的重要性。随着新媒体崛起及媒体融合深入，传播方式多样化导致版权侵权问题日益严重。深度挖掘多模态内容特征信息，可实现对新媒体内容全方位版权保护。【方法】文章针对新媒体领域中版权保护生态现状，分析目前新媒体版权保护面临的主要问题，介绍基于数字水印、多模态内容指纹和区块链技术打造的“南方版专 +”版权保护平台。【结果】通过平台应用效果，可验证该平台对新媒体中多模态内容实现快速确权，降低版权保护成本。【结论】通过“南方版专 +”版权保护平台，可以有效解决新媒体中因侵权方式多样化带来的确权难、维权成本高等问题。

关键词：新媒体；数字版权；多模态内容指纹；区块链技术；数字水印　**中图分类号：**G236　**文献标识码：**A
文章编号：1671-0134（2023）06-074-06　**DOI：**10.19483/j.cnki.11-4653/n.2023.06.015
本文著录格式：王梦环，罗小龙，李梓华．基于多模态特征挖掘的新媒体版权保护方法研究与应用 [J]. 中国传媒科技，2023（06）：74-78，86.

导语

据国家版权局发布的最新版《中国网络版权产业发展报告（2020）》中指出，2020 年中国网络版权产业市场规模首次突破一万亿元 [1]，较 2019 年增长 23.6%，用户付费规模接近 5659.2 亿元，同比增长 27.3%。网络新闻媒体和网络游戏仍占据主导地位，而短视频、直播等新兴业态迅速崛起，推动网络版权内容走向“视频化”和“价值向”。

南方报业作为以广东省委机关报《南方日报》为龙头的传媒集团，经过几十年的数字化变革，积累了丰富的知识版权资源。面对移动互联网时代，南方报业传媒集团不断创新图文、音视频等全媒体原创内容，逐步从传统媒体转向微信、微博、客户端等为代表的新媒体媒介，版权保护工作随着传播内容和传播方式的变化面临新的挑战。因此，南方报业传媒集团研究并探索了一种适用于新媒体多模态数字内容的版权保护方法，打造集确权、监测、存证、维权为一体的一站式版权服务平台“南方版专 +”，提供全链条版权保护解决方案，为各类媒体单位在新媒体环境下的版权保护提供参考。

1. 新媒体领域数字版权保护面临的主要问题

当前社会已经发展成为以手机等移动设备为主要载体的信息传播时代，新媒体内容具有传播速度快、移动性强、互动性好、个性化足等优势，但新媒体平台上盗版、抄袭、洗稿等层出不穷的侵权方式让内容生产商利益受损，创作积极性备受打击。虽然国家非常重视，出台多种政策和法规，但面对新媒体技术的迅猛发展，以及侵权平台多样化、侵权手段日益隐蔽等局面，新媒体作品内容容易修改、取证环境复杂、证据容易销毁，造成权利人确权难、取证难，由于缺乏有效的技术手段，原创内容的版权保护很难从根本上解决问题。

同时，在线下维权过程中，整个公证流程繁琐，且存在侵权方利用纸质公证文书造假的方式进行干扰等问题，导致在维权过程中需要消耗大量的人力、物力且效率低下，因此亟须一种高效且公正的维权方式来解决维权难的问题。

2.“南方版专 +”多模态版权保护方法研究

2.1　总体框架设计

“南方版专 +”平台能够针对原创的文本、图片、视频等内容从创作、发布、传播等环节进行全链条的跟踪记录，将版权作品的每个变动记录在区块链上。

作者在完成原创创作后，平台依据原创内容属性，自动化进行内容加密。对于文本内容，基于指纹哈希技术为其生成内容指纹信息（也称为内容 *DNA*）；对于图片，基于图像变换自适应数字水印技术进行加密嵌入和提取；对于视频内容，基于视频特征与时间序列信息进行内容指纹信息生成。

在对原创内容生成内容指纹信息后，系统使用私钥签名将加密后的数据发送给区块链，并将数据版权的明文信息留存本地。区块链的记链接口在收到明文信息后，智能合约首先对记链信息验证签名，签名通过则将记链信息记入账本并发送其他节点进行账本同步。[2] 此时存放在区块链系统中的信息能完整反映作品的创作和传播过程中的存证序列，能为版权出现争议时提供更多可以检索和验证的关键信息，推动版权保护的模式从以作品结果为中心转为以创作过程为中心。

在侵权检测过程中，依据需要对比的内容模态，选择对应的特征转换操作。结合内容相似度比对计算和数字水印提取及比对这两种侵权检测手段同时进行，主要为解决原创内容被侵权方通过技术手段多重转换的问题，提高检测能力的鲁棒性。对检测出来的疑似侵权结果，会由人工进行介入来二次确认，经由人工确认并确定的侵权内容会将其信息进行上链，并依靠区块链的不可篡改、可溯源等特性进行固证来用于辅助后续维权工作。

系统架构图如下图所示。

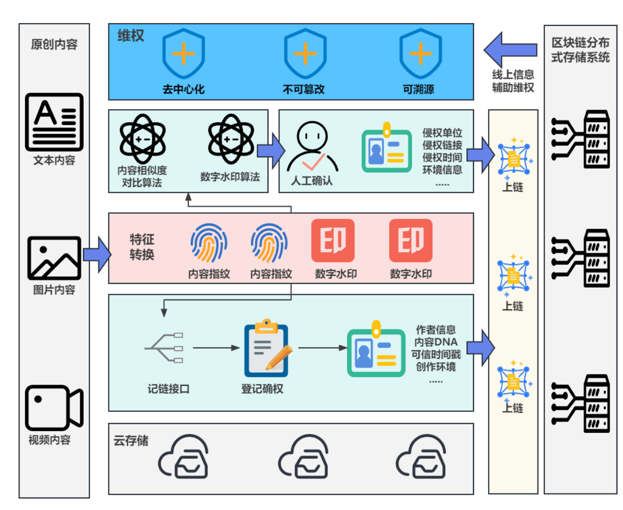


图 1 “南方版专 +”系统架构图

2.2 关键技术介绍

2.2.1 基于 MinHash 的文本内容指纹生成技术

由于文本的长度、大小均不一致，对全文的所有信息直接上链显然不现实，平台采用内容指纹算法对文本的内容、摘要等信息进行计算，得到文本的指纹信息，该指纹信息一般为一个长度较小的字符串。即：通过内容指纹映射函数，将原始复杂且冗长的文本信息映射为一个短小的字符串 *DNA* 信息，因此可以大大缩短上链时间，提升上链效率。文本指纹提取的原理如下。

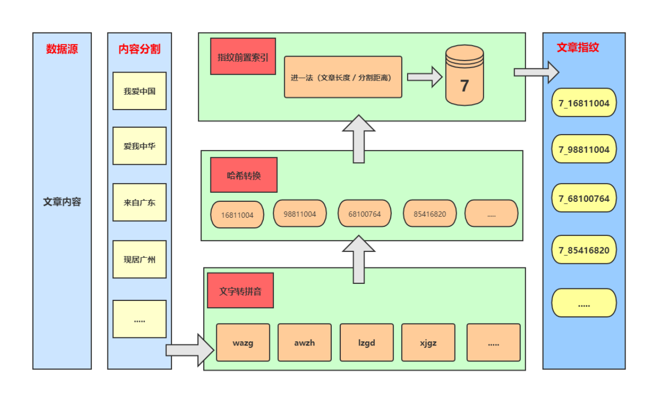


图 2 文本内容指纹生成流程图

首先将文章内容进行 x 等距切分成 n 段（$x \in [1,100]$），并且对每个分段的内容进行文字转拼音，取每个词语拼音首个字母作为指纹转换特征（例如：我 –> *wo*–> *w*）；其次通过 MinHash 算法（最小哈希是适用于大规模高纬度数据的一种快速最近邻查找算法，其实质是基于一种假设，相似度很高的两个数据映射成同一个 hash 值的概率较大，而相似度很低的两个数据则很难映射成同一个 hash 值 [3]）对指纹转换特征进行哈希转换得到指纹信息，如：我爱中国 –> *wazg* –> *hash*(*wazg*) –> 16811004，然后根据文章的内容长度给指纹信息定义前置索引，就可以得到文章完整的指纹信息。例如：某一篇文章的内容长度是 26 个字符，拟将切分距离 x 设为 4，则可以将该文章等距切分成 7 段文本，最终生成的 7 个带前置索引的指纹信息为该文章的指纹信息。

2.2.2 基于 CRBM 的视频内容指纹生成技术

在视频内容指纹技术出现之前，对视频相关性的比对分析主要依赖于视频文件的哈希值进行比对。虽然此方式对验证文件一致性具有计算快、体积小等优势，但在视频相关性的比对分析中仍存在明显不足。主要原因是基于哈希值的比对分析只在整个视频内容或视频内容中的某些片段完全相同时才有效。在实际工作中，即使是同一视频也会存在因不同的视频编码或数字处理版本而导致视频文件的哈希值不同的情况。

因此采用基于深度学习算法的视频内容指纹生成技术将视频特征与时间序列信息相结合进行特征计算，来解决哈希值比对在视频相关性计算中的不足问题。视频内容指纹生成技术是一种使用特征提取算法对视频内容进行识别、提取、压缩后生成一个比原视频内容小很多的视频指纹的方法，该指纹可标记视频唯一身份，不会因视频内容的格式转换、剪辑拼接、压缩

旋转等变换而发生变化。通过对两个视频的指纹进行比对分析，可快速对比待检测视频与原视频之间的相关性，从而达到版权保护的目的。

在进行视频内容指纹生成时，须先对目标视频进行逐帧解码，获取该视频的图像序列。因相邻像素和连续帧之间的相关性反映了每一帧的局部结构及其时间动态，此种相关性为具备较强辨别力和稳定性的视觉特征。为捕捉和变现这种抽象的视觉特征，将条件受限玻尔兹曼机器[4]（Conditional Restricted Boltzmann Machine，CRBM）作为构建深度特征学习网络的关键部件之一。CRBM 使用如下能量函数来定义过去 m 帧上可见和隐藏层的联合概率：

$$E(v_t,h_t)=\frac{1}{2}\left\|v_t-(\sum_{k=1}^{m}A_k v_{t-k}+a)\right\|_2^2-(\sum_{k=1}^{m}B_k v_{t-k}+b)^T h_t-v_t^T W h_t$$

其中 W 、A_k 和 B_k 为权重矩阵，a 和 b 为偏移参数，用于平衡可见层与隐藏层之间的关系。CRBM 的结构示意图如下：

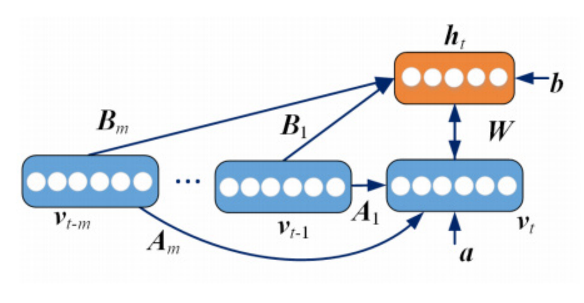


图 3　CRBM 结构示意图

通过训练参数 $\{W,a,b,A_k,B_k\}$ 来求解最小化损失函数，函数表达式如下：

$$L_{CRBM}=-\ln\sum_{h_t}\exp(-E(v_t,h_t))+\ln\sum_{h_t}\sum_{v_t}(-E(v_t,h_t))$$

通过最小化损失函数，从而求出可见层（多帧图像）对应的隐藏层，即视频指纹。

2.2.3　基于内容指纹的相似度比对技术

2.2.3.1　基于文本内容指纹的相似度对比技术

基于文本内容指纹的相似度对比技术是根据指纹信息与内容唯一 D 进行关联分析，其中指纹信息与内容 D 是一对多的关联关系，关联的前提条件是：假如稿件 A 中存在指纹 $F1$ 信息，则表示指纹 $F1$ 与稿件 A 关联成功，最终将所有指纹信息与内容 D 形成关联图谱存储在数据库中，实现快速、准确的文本内容相似度对比效果。

对比过程中如图 4 所示，首先将文本内容 A 进行指纹生成，得到指纹列表 F，将指纹列表 F 中所有指纹 $\{f_1,f_2,......,f_n\}$ 作为查询条件去查询指纹数据库，查询结果则返回若干个内容 D 信息 $\{i_1,i_2,......,i_n\}$，$\{i_1,i_2,......,i_n\}$ 则表示指纹 f 的查询结果，即 $\{i_1,i_2,......,i_n\}$ 这些稿件内容都共同拥有 f 指纹信息，最终将内容 D 列表进行统计分析，每个 D 在列表中每出现一次即代表指纹命中一次，以此类推，将得到所有内容 D 的指纹命中次数表 $\{i_1:x_1,i_2:x_2,......,i_n:x_n\}$，$x$ 则代表命中次数，命中次数越高则表示与文本内容 A 越相似。

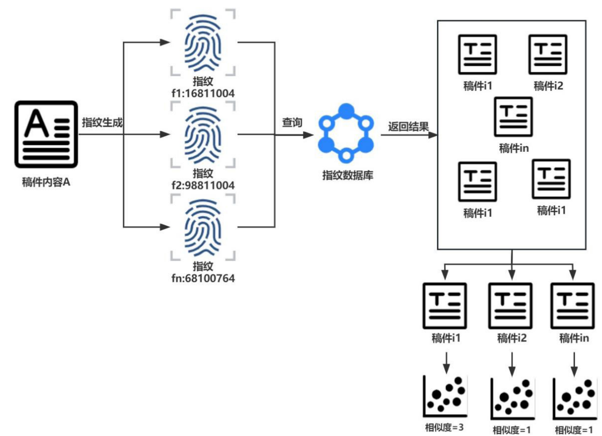


图 4　文本内容相似度比对流程图

2.2.3.2　基于视频内容指纹的相似度对比技术

基于视频内容指纹的相似度对比技术是根据指纹时间线进行关联分析，其中视频内容下有 n 个指纹信息，每个指纹信息都有记录在视频中出现的时间线，根据指纹时间线进行对比分析[5]，即可快速、准确地找到目标指纹与库存指纹中的相似视频片段。

对比的过程如图 5 所示，视频 A 作为相似度对比的主体，将视频 A 进行指纹生成，得到指纹和指纹时间线，在指纹库中得到视频 B 的指纹和指纹时间线，将指纹 f_{a1} 进行时间线固定，对视频 B 中的指纹所在的时间线上进行逐帧滑动，在滑动过程中计算 A 与 B 之间指纹序列的重叠次数，从中得到疑似相似的片段，再结合灰度信息进行相似度算分，最终得到视频 A 与视频 B 的相似度得分。

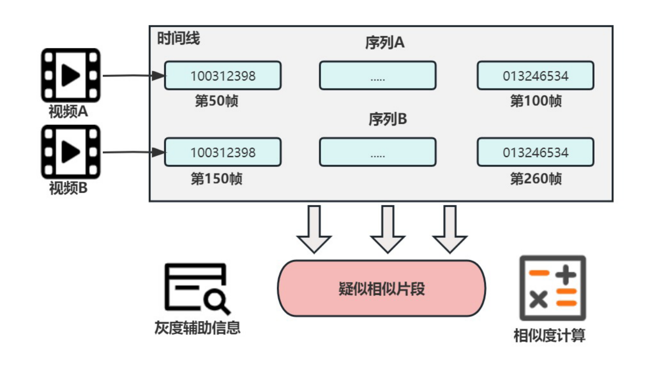


图 5　视频内容相似度比对流程图

2.2.4 数字水印嵌入与提取技术

数字水印技术作为一种典型的信息隐藏技术，把标识信息（如：商标、版权声明、图章、电子签名等标识性内容）直接嵌入数字载体当中（如：多媒体、文档、软件等），嵌入原则上不影响数字载体的使用价值，且数字水印也应该是不易被探知、篡改和擦除的[6]，以此来实现防伪溯源、版权保护的目的。

数字水印的架构如图6所示，主要基于 *DCT* 图像变换自适应数字水印技术进行嵌入和提取：水印嵌入阶段主要将原始数据与加密后的水印通过水印嵌入算法加工，获取到含隐藏水印的数字载体；水印提取阶段主要通过水印提取算法来识别数据中是否含有水印，提取数据中的水印内容，进而定位到数据的具体来源。

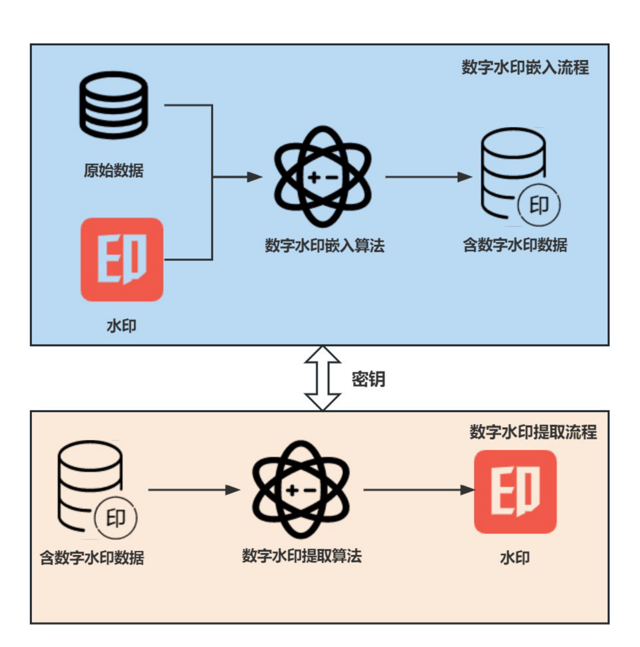


图 6　数字水印嵌入与提取流程图

在采用 *DCT* 图像变换自适应数字水印技术[5]将水印信号嵌入到原始图像载体过程中，主要分为水印图像预处理、宿主图像变换、数字水印嵌入、数字水印提取。

在水印图像预处理的过程中，为了减弱图像数据的相关性，将水印图像进行二值化处理，有像素的位置为1，否则为0，并对水印进行图像分块，图像块的数目与原始图像保持一致，设为 $W=\{W(i,j)|0\le i<M,0\le j<N,W(i,j)\in\{0,1\}\}$。从而保证了视觉对嵌入水印的不敏感性和水印信息的安全性。

宿主图像变换则是对 N×N 大小的256灰度的宿主图像进行 8×8 的不重叠像素分块，将宿主图像从空间域转换到 *DCT* 变换域，以 *Zig-Zag* 方式对 *DCT* 变换域的图像频率系数进行重新排列，得到一组一维向量系数 $D=\{d_1,d_2,......,d_n\}$，并从中选出中频系数 $D_l=\{D_l+1,D_l+2,......,D_l+M\}$。

基于数字水印预处理下得到的数字水印 W 服从标准正态分布，数字水印嵌入的过程中采用 W 随机序列对中频系数进行修改，公式如下：

如果 $w_i=0$，则 $D_i(k)'=\frac{1}{5}\sum_{l=k-2}^{k+2}D_i(l)-Q_i$

如果 $w_i=1$，则 $D_i(k)'=\frac{1}{5}\sum_{l=k-2}^{k+2}D_i(l)+Q_i$

其中 Q_i 为水印强度调节因子，受图像块 i 的掩蔽特性值 $S_m(i)$ 决定，a,b 分别为 Q_i 和参数 $S_m(i)$ 线型比例参数。计算方法如下：

$$Q_i=a+b\times S_m(i)$$

在水印提取时，则选取相同的 *DCT* 系数，进行上述逆流程操作，即可提取出数字水印。

2.2.5 区块链链上 + 链下存储

区块链技术是一种去中心化的分布式数据库技术，由多个节点共同维护一个不断增长的数据链。区块链存储技术可分为链上存储和链下存储两种。其中链上存储适用于存储关键信息，以确保其安全性和不可篡改性；链下存储则更适用于存储大量数据，以降低成本和保护隐私。

在实际应用中，由于区块链的存储空间有限，链上存储大量数据会导致存储成本较高。此外，链上数据的公开性存在隐私泄露的问题。针对此问题采取链上 + 链下的设计模式[7]：采用明确授权与选择性披露的存储策略，用户将身份标识的信息公布到链上，涉及隐私的数据存储在链下，能够有效保护数据的隐私安全，同时也降低数据的存储成本。

链上存储的过程中，将数据直接存储在区块链的区块中，每个区块都包含一组经过加密的数据，当一组加密数据被确认后，就会被打包进一个区块，然后添加到区块链上，一旦数据被写入区块链，就无法进行修改和删除，保证了数据的完整性和安全性，写入过程中采用分布式存储在各个参与者的节点上，降低单点故障风险，同时能够被所有参与者查看，确保数据的可追溯性和透明度。

链下存储的过程中，涉及繁琐运算的检验与核算工作拆解到链下，将无须参与共识、数据量大、计算繁琐的数据存储在链下的数据库或分布式文件系统中。

当特定链下数据需要进行多方共识时，可通过链下本地化分析处理后，再与链上进行交互，通过接口发送交易完成上链。

2.3 “南方版专 +”特点

“南方版专 +”是由南方报业传媒集团针对目前版权市场行业现状和媒体维权痛点打造的集确权、监测、存证、维权于一体的一站式版权服务平台。[8] 平台致力于为各类媒体提供全方位的版权运维解决方案，助力媒体版权保护与版权价值变现。平台具有以下特点。



图 7 “南方版专 +”平台简介图

2.3.1 创作即确权

“南方版专 +”从区块链资产的角度切入数字版权，利用区块链技术中的不可篡改性，将内容作者、原创时间和创作内容三位一体进行确权。平台对发布后的内容、作者真实姓名、创作环境、创作时间等信息通过现代密码学算法进行数字签名和加密，生成内容的数字指纹，加上可信的时间戳后进行上链，在得到其他节点的确认后，即完成内容确权的动作，从而保证数据的可信及不可篡改，实现“创作即确权”。

2.3.2 覆盖全网，智能侵权监测

“南方版专 +”涵盖纸媒、网站、客户端、微信、微博、论坛等多平台内容数据库，监测覆盖报纸 1400+，网站 18000+，移动 App1300+，微信公众号 13 万 +，微博 5 万 +。利用大数据智能分析技术，对用户发布的原创作品进行 7*24 小时不间断全网监控并及时反馈监测结果，让侵权行为无所遁形。同时，平台可针对用户需求进行个性化检测，如版权方可根据实际需要选择单张监测或接入图片库作批量监测，随后可实现对图片单独维权，以体现其应有的版权价值。

2.3.3 一键取证，永久固证

针对监测发现的侵权记录，平台支持一键申请证据保全，实时固化侵权网页内容并将其封存在服务器上，在经过加密处理后平台将证据同步至公证处服务器上，形成具备公证效力的电子证据。根据维权用途，对存证数据在线向公证处申请出具《电子数据公证保管证书》或《电子证据保全公证书》，节约固证成本，确保证据效力。固证后的证据可通过电子证据平台直接递交互联网法院的区块链平台——网通法链，极大提高维权效率，同时降低维权成本。

2.3.4 线上线下一体化维权

“南方版专 +”还提供律师函发送、侵权调解、提起诉讼等专业律师维权服务，让入驻单位轻松维权。作为一家由媒体开发的版权服务平台，南方版专 + 比任何一家商业技术平台更了解媒体需求和媒体生态，为平台用户提供多元化的版权服务，共同推动版权生态的健康良性发展。



图 8 “南方版专 +”维权流程图

2.3.5 应用与成效

“南方版专 +”平台独有集确权、监控、取证及线上开庭于一体并直连广州互联网法院立案系统，目前已在南方报业传媒集团全面应用，目前已立案的作品超万件。除南方报业传媒集团报网端自用外，目前平台已吸引安徽日报报业集团、海南日报报业集团以集团名义携旗下全部媒体整体入驻，同时广东省内的南方财经传媒集团下属《21 世纪经济报道》，广东省出版集团下属《时代周报》；广东省外的《中国经营报》《上海证券报》等财经类媒体也已成功入驻，在全国财经媒体联盟形成了一定影响力。以《21 世纪经济报道》的入驻为例，3 年来，平台为其代理维权平台超过 60 家，目前达成和解的超过 20 家，涉及合同金额超过 700 万。

结语

随着深度学习技术和区块链技术的高速发展，“南方版专 +”也将随着技术生态的不断优化和应

（下转第86页）